APPLICATION OF DEEP LEARNING FOR THE QUANTITATIVE ASSESSMENT OF BONE MARROW LESIONS (BMLs)

*Brigham and Women’s Hospital, Harvard Medical School

INTRODUCTION: MRI can be used to assess several structural changes related to knee OA (KOA). Software image analysis methods provide objective fully quantitative measurements of these changes. However, such methods generally require a human reader and can be time consuming if the number of images in a study is large. Deep leaning (DL), a form of statistical machine learning, offers the potential for increased or full automation, therefore reducing reader time substantially.

OBJECTIVE: The goal of our study was to validate a DL method for segmenting (outlining) bone marrow lesions (BMLs) on knee MRI data sets.

METHODS: We used the baseline and 24 month scans from the 600 subjects included in the FNIH Study, a nested case/control sub study of the OAI. Subjects in the sub-study were selected from four groups based on medial radiographically-assessed joint space loss (JSL) and pain progression. JSL progression was defined as medial tibiofemoral radiographic JSL≥0.7mm from baseline to 24, 36, or 48 months. Pain progression was defined as persistent increase on the WOMAC pain scale from 24 to 60 months. For the DL method, the reader was required to indicate the location of each BML, a step that requires less than 30 seconds of reader time per scan. The readings were performed on sagittal turbo spin echo fat-suppressed (TSE FS) MRI scans at the baseline and 24-month time points. The reader was fully blinded.

BMLs segmented using a previously validated semi-automated (SA) method[1] were used to train the DL algorithm using a cross-validation technique. Once trained, the DL software was applied to the full data set to calculate volume and comparisons were made to the results from the SA method. Dice similarity coefficients (DSC) and Pearson’s R² are reported. For clinical validation, we used logistic regression with odds ratios (OR) and 95% confidence intervals (CI) as metrics to assess the association of baseline to 24 mo. change in BML volume with progression status. Each model was adjusted for age, sex, BMI, and race. ORs were calculated per 1 standard deviation loss of cartilage from baseline to 24 months. In Figure 1 we show a representative example of BMLs segmented with SA and DL.

RESULTS: The average DSC was 0.85 and Figure 2 is a plot comparing the SA to the DL methods. The Pearson’s R² was 0.99 and the slope was 0.85 indicating that the DL systematically underestimates the volume. The OR’s (CI) for the SA and DL methods were 0.81 (0.67, 0.96) and 0.80 (0.68, 0.97) respectively.

CONCLUSIONS: The results suggest that the SA and LD methods are nearly equivalent for segmenting BMLs for KOA. The DL approach can produce a highly accurate method to assess BML volume that requires offers substantial time savings over the SA method. The 0.85 shift can easily be corrected. Future work will investigate a fully automated method, which does not require a reader to initially locate BMLs.

SPONSOR: BWH Brigham Research Institute

DICLOSURE STATEMENT: None

CORRESPONDENCE ADDRESS: jduryea@bwh.harvard.edu